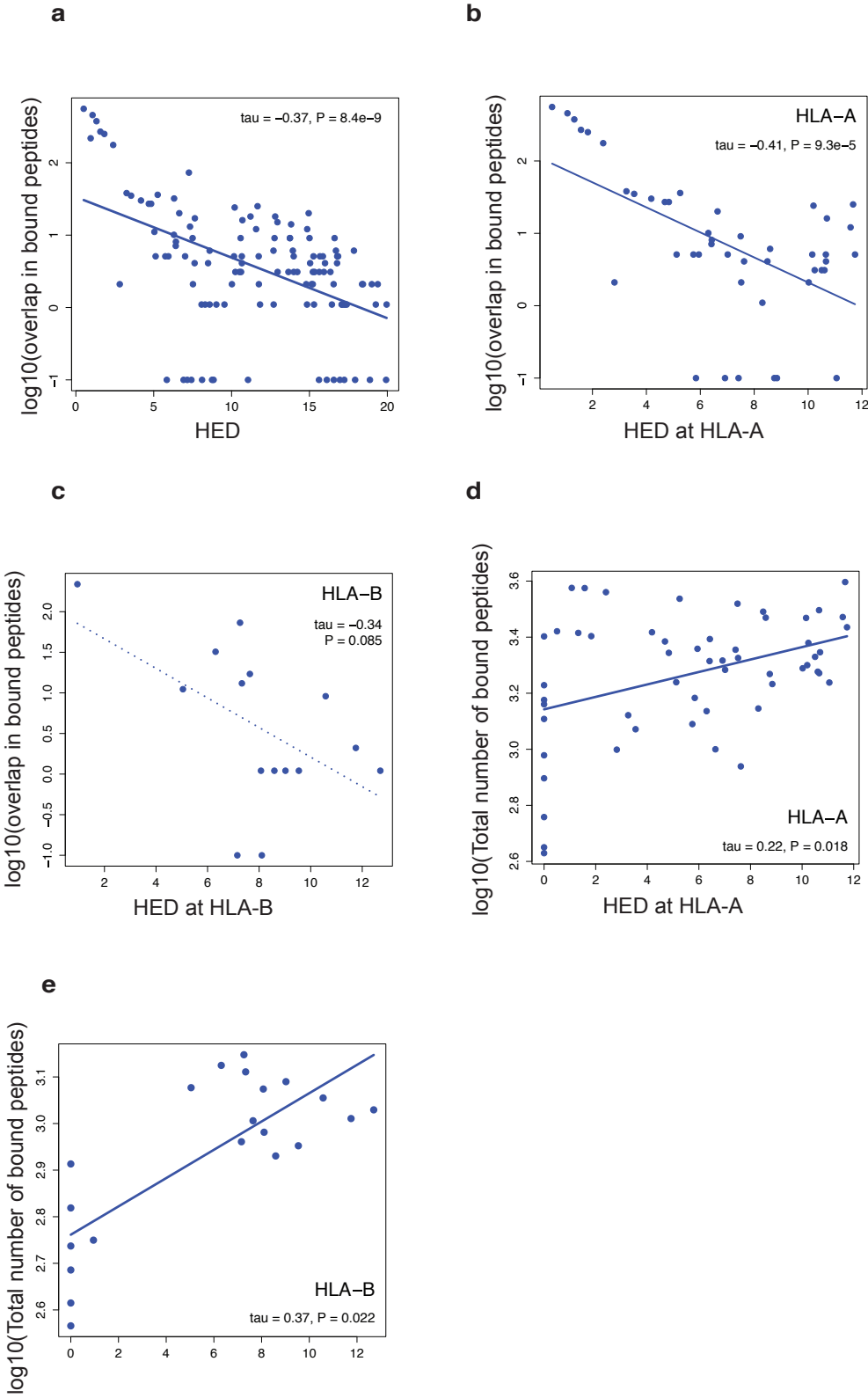


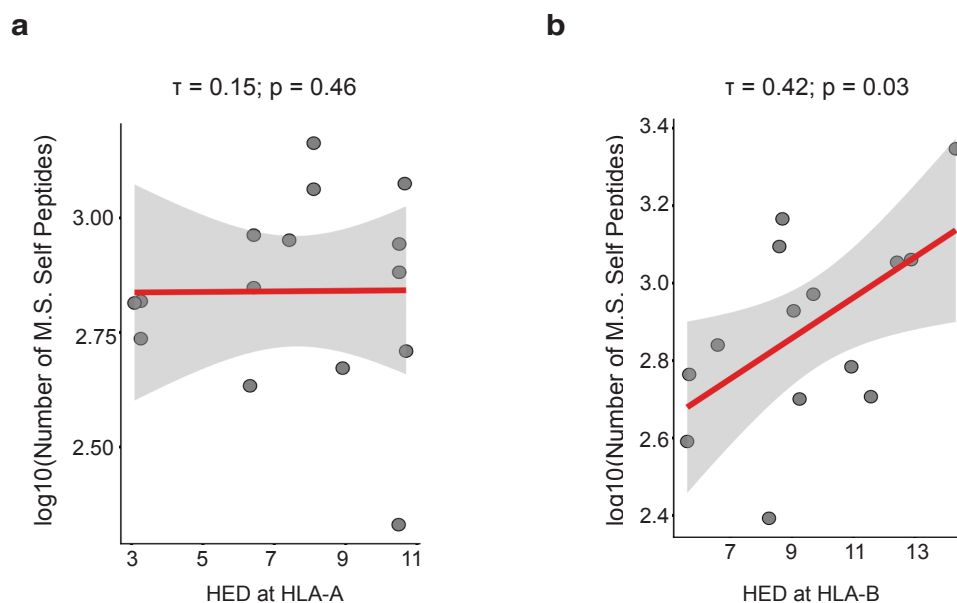
Supplementary Figure 1



**Supplementary Figure 1 | Validation of Grantham distance score between alleles using the mass spectrometry peptidomes derived from mono-allelic cells by Abelin et al.**

**a**, The Grantham distance score, used here to estimate HLA evolutionary divergence (HED) between *HLA* alleles, correlates negatively with the overlap of peptides bound by any two *HLA* alleles from the dataset of Abelin et al., which contains naturally eluted peptide repertoires from mono-allelic cell lines of 16 different *HLA*-I alleles (representing 120 possible allele pairs). **b**, Same as **a**, for *HLA*-A alleles alone ( $n = 45$  allele pairs). **c**, Same as **a**, for *HLA*-B alleles alone ( $n = 15$  allele pairs). Blue line indicates line of best fit. **d**, HED at *HLA*-A is positively correlated with the total number of peptides bound to each pair of *HLA*-A alleles ( $n = 55$  allele pairs). **e**, Same as **d**, for *HLA*-B alleles ( $n = 21$  allele pairs). All p-values calculated were calculated using two-sided Kendall's rank correlation.

## Supplementary Figure 2



### Supplementary Figure 2 | Association of HLA-I evolutionary divergence at *HLA-A* and *HLA-B* with diversity of the self immunopeptidome generated by mass spectrometry from Pearson et al.

**a**, Correlation of HED at *HLA-A* with number of unique naturally processed self peptides bound to alleles of each *HLA-A* genotype from patients with metastatic melanoma patients heterozygous at HLA-A ( $n = 15$ ) from Pearson *et al*;  $P = 0.46$ ; two-sided Kendall's rank correlation. Data recapitulate results derived from computational peptide-HLA binding predictions shown in Extended Data Fig. 10g. **b**, Correlation of HED at *HLA-B* with number of unique naturally processed self peptides bound to alleles of each heterozygous *HLA-B* genotype ( $n = 15$ );  $P = 0.03$ ; two-sided Kendall's rank correlation. Data recapitulate results derived from computational peptide-HLA binding predictions shown in Extended Data Fig. 10h. Red line indicates line of best fit.